

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

S 00 P 1087 Y S 00
Priority Paper
2/13/01
MOH
Jc658 U.S. PTO
09/661878
09/14/00

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日

Date of Application:

1999年 9月14日

出願番号

Application Number:

平成11年特許願第261262号

出願人

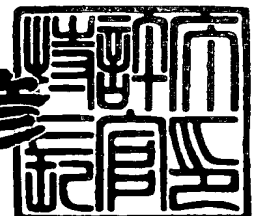
Applicant(s):

ソニー株式会社

2000年 6月29日

特許庁長官
Commissioner,
Patent Office

近藤 隆彦



出証番号 出証特2000-3050041

【書類名】 特許願

【整理番号】 9900066903

【提出日】 平成11年 9月14日

【あて先】 特許庁長官殿

【国際特許分類】 H04N 7/133

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社
内

【氏名】 カーティス ユーバンクス

【特許出願人】

【識別番号】 000002185

【氏名又は名称】 ソニー株式会社

【代表者】 出井 伸之

【代理人】

【識別番号】 100067736

【弁理士】

【氏名又は名称】 小池 晃

【選任した代理人】

【識別番号】 100086335

【弁理士】

【氏名又は名称】 田村 榮一

【選任した代理人】

【識別番号】 100096677

【弁理士】

【氏名又は名称】 伊賀 誠司

【手数料の表示】

【予納台帳番号】 019530

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9707387

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 信号処理装置及び方法並びに記録媒体

【特許請求の範囲】

【請求項 1】 入力画像信号の形式を特定する形式特定手段と、
上記入力画像信号に対する編集操作が入力される入力手段と、
上記入力画像信号を上記入力手段に入力された編集操作にしたがって編集処理する、上記形式に対応する複数の編集処理手段と、
上記編集処理手段にて編集された上記入力画像信号を出力ストリームに変換する変換手段と、
上記形式特定手段にて特定した上記入力画像信号の形式に応じて、当該形式に対応する上記編集処理手段で編集処理を行うように制御する制御手段と
を有することを特徴とする信号処理装置。

【請求項 2】 上記形式とは、上記入力画像信号のフォーマット又はフォーマットのクラスであることを特徴とする請求項 1 記載の信号処理装置。

【請求項 3】 上記編集操作は、上記入力画像信号のカット、コピー、ペースト、リサイズ、時間的な移動、テキスト又は画像の重畳、トランジションを含むことを特徴とする請求項 1 記載の信号処理装置。

【請求項 4】 上記編集処理手段における、特定の形式に対する編集処理を置換、削除又は置換する編集処理管理手段を有することを特徴とする請求項 1 記載の信号処理装置。

【請求項 5】 上記入力画像信号を共通フォーマットに変換するフォーマット変換手段を有し、上記編集処理手段は、上記フォーマット変換手段で共通フォーマットに変換された画像信号に対して編集処理を施すことを特徴とする請求項 1 記載の信号処理装置。

【請求項 6】 入力画像信号の形式を特定し、
上記入力画像信号に対する編集操作が入力され、
上記入力画像信号を入力された編集操作にしたがって編集し、
編集された上記入力画像信号を出力ストリームに変換し、
特定した上記入力画像信号の形式に応じて、当該形式に対応する編集を行うよ

うに制御すること

を特徴とする信号処理方法。

【請求項 7】 上記形式とは、上記入力画像信号のフォーマット又はフォーマットのクラスであることを特徴とする請求項 6 記載の信号処理方法。

【請求項 8】 上記入力画像信号を共通フォーマットに変換することを特徴とする請求項 6 記載の信号処理方法。

【請求項 9】 画像信号を処理するプログラムが記録された記録媒体であって、上記プログラムは、

入力画像信号の形式を特定し、

上記入力画像信号に対する編集操作が入力され、

上記入力画像信号を入力された編集操作にしたがって編集し、

編集された上記入力画像信号を出力ストリームに変換し、

特定した上記入力画像信号の形式に応じて、当該形式に対応する編集を行うように制御すること

を特徴とする記録媒体。

【請求項 10】 上記形式とは、上記入力画像信号のフォーマット又はフォーマットのクラスであることを特徴とする請求項 9 記載の記録媒体。

【請求項 11】 上記入力画像信号を共通フォーマットに変換することを特徴とする請求項 9 記載の記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、画像信号を処理する信号処理装置及び方法並びに画像信号を処理するプログラムが記録された記録媒体に関する。

【0002】

【従来の技術】

従来、非線形ビデオ編集装置と称される画像編集装置が提供されている。この非線形ビデオ編集装置は、編集の際の一時記録媒体としてハードディスクの如きディスク状記録媒体を備えるので、ランダムアクセスが可能である。したがって

、シーケンシャルではない非線形な編集が可能である。かかる非線形ビデオ編集装置においては、所望の映像素材の迅速な読み出しと、一連の映像素材に対する素材の挿入及び削除を容易に行うことができる。

【 0 0 0 3 】

【発明が解決しようとする課題】

しかし、従来の非線形ビデオ編集装置は、ユーザーに単一の画像データのフォーマットを使用するように制限するか、又は複数のフォーマットの使用を可能とさせるが、特定の効果又は全部の効果（例えば、トランジション効果）の画像処理を行う場合に、画像データを共通フォーマットに変換するように制限していた。

【 0 0 0 4 】

前者の場合には、単一フォーマットを使用するように制限すると、ユーザがフォーマットを自由に選択する自由を妨げることになり、ユーザへ負担を掛けることになる。

【 0 0 0 5 】

また、後者の場合には、共通フォーマットに変換する処理は、顕著な処理上のオーバーヘッドが導入されたり、高価なハードウェアが必要とされたり、ビデオ画像の品質の劣化が生じたりする。

【 0 0 0 6 】

特に、いわゆる M P E G (Moving Picture Experts Group)、いわゆる J P E G (Joint Photographic Coding Experts Group)、いわゆる H. 2 6 1、いわゆる D V (Digital Video) 及びその他の通常用いられるデジタル画像フォーマットのように、離散コサイン変換 (Discrete Cosine Transformation; DCT) に基づく画像フォーマットは、圧縮処理の一部として D C T 係数の量子化に依存している。

【 0 0 0 7 】

汎用エディタの普通の技術では、このような画像フォーマットを画素 (picture element; pixel) 表示に復元し、空間領域内で画像処理を実行し、その画像を再び圧縮符号化する。しかし、符号化の際における D C T 係数の再量子化と整数

化は、画像歪みをもたらし、画像歪みは世代を重ねるごとに累積してゆく。

【0008】

本発明は、上述の実情に鑑みて提案されるものであって、画像信号の編集を高速に、画質を劣化させないように行える信号処理装置及び方法並びにかかる編集のプログラムが記録された記録媒体に関する。

【0009】

【課題を解決するための手段】

上述の課題を解決するために、本発明に係る信号処理装置は、入力画像信号の形式を特定する形式特定手段と、上記入力画像信号に対する編集操作を入力される入力手段と、上記入力信号を上記入力手段に入力された編集操作にしたがって編集処理する、上記形式に対応する複数の編集処理手段と、上記編集処理手段にて編集された上記入力画像信号を出力ストリームに変換する変換手段と、上記形式特定手段にて特定した上記入力信号の形式に応じて、当該形式に対応する上記編集処理手段で編集処理を行うように制御する制御手段とを有するものである。

【0010】

本発明に係る信号処理装置は、入力画像信号の形式を特定し、上記入力画像信号に対する編集操作を入力され、上記入力画像信号を入力された編集操作にしたがって編集し、編集された上記入力画像信号を出力ストリームに変換し、特定した上記入力画像信号の形式に応じて、当該形式に対応する編集を行うように制御するものである。

【0011】

本発明に係る記録媒体は、画像信号を処理するプログラムが記録された記録媒体であって、上記プログラムは、入力画像信号の形式を特定し、上記入力画像信号に対する編集操作を入力され、上記入力画像信号を入力された編集操作にしたがって編集し、編集された上記入力画像信号を出力ストリームに変換し、特定した上記入力画像信号の形式に応じて、当該形式に対応する編集を行うように制御するものである。

【 0 0 1 2 】

【発明の実施の形態】

以下、本発明の実施の形態について、図面を参照して説明する。まず、本発明の実施の形態としての画像編集装置の構成を、図 1 を参照して説明する。本実施の形態の画像編集装置は、複数の画像データのフォーマットを取り扱うことができるものであり、フォーマットに応じて拡張が可能である。

【 0 0 1 3 】

この画像編集装置は、特定のフォーマット又はフォーマットのクラス、すなわち、MPEGのようなDCTに基づいた周波数領域及び圧縮領域での処理のための高度に最適化された処理に対応する複数のビデオフォーマットモジュールを備える。この画像編集装置においては、圧縮領域と周波数領域における処理技術を用いることによって、多くの画像処理をより速く、少ない劣化にて行うことができる。

【 0 0 1 4 】

図 1 に示すように、本実施の形態の画像編集装置の好適なハードウェア構成は、画像編集装置の全体を制御するCPU 1 と、高速のメモリであるキャッシュ 2 と、システムに対する制御を行うシステム制御部 3 と、データを記憶するメモリ 4 とを有している。

【 0 0 1 5 】

CPU 1 は、この画像編集装置の全体を制御すると共に、各種の処理を実行する。キャッシュ 2 は、高速メモリであって、頻繁にアクセスされるメモリ 4 の部分を含み、これによってメモリ 4 へのアクセス速度を増加させる。システム制御部 3 は、CPU 1、キャッシュ 2、メモリ 4 を含むシステムの資源を制御するために使用される。

【 0 0 1 6 】

メモリ 4 にはデータが格納される。メモリ 4 は、CPU 1 及びシステム制御部 3 からのコマンドに基づいて、データを書き込まれたり読み出されたりすることができる。本実施の形態においては、メモリ 4 にはDRAMが使用される。

【0017】

CPU 1、キャッシュ 2、システム制御部 3 及びメモリ 4 は、内部バス 1 1 にて相互に接続されている。内部バス 1 1 は、これらの間の高速通信手段として作用する。

【0018】

また、画像編集装置は、大容量の記録装置であるハードディスク 6 と、ハードディスク 6 を制御するディスク制御部 5 と、ポインティングデバイスであるマウス 8 と、マウス 8 を制御するマウス制御部 7 と、文字や画像を表示するモニタ 1 0 と、モニタ 1 0 の表示を制御する画像制御部 9 とを有している。

【0019】

ハードディスク 6 は、固定ディスクを用いた大容量の記録装置である。ディスク制御部 5 は、ハードディスク 6 の指定された領域に対するデータの書き込みと読み出しを制御する。

【0020】

画像編集装置においては、ハードディスク 6 を用いて画像データにランダムアクセスすることが可能であり、画像データに対するシーケンシャルでない非線形編集を行うことができる。

【0021】

マウス 8 は、その動きに基づいた位置データが得られるポインティングデバイスである。また、マウス 8 は、ボタンを押圧されることによってもデータが入力される。CPU 1 は、マウス 8 から送られたデータに基づいて、モニタ 1 0 上にカーソルを表示して、ユーザーとの間で双方向的に制御を行う。

【0022】

モニタ 1 0 は、画像を表示するものであり、例えば CRT が利用される。モニタ 1 0 は、CPU 1 によって作成された状態に基づいて、時間変化する画像を表示する。画像制御部 9 は、CPU 1 から描画コマンド又は他のコマンドを受け取って処理し、画像信号に変換してモニタ 1 0 に送る。

【0023】

システム制御部 3、ディスク制御部 5、マウス制御部 7 及び画像制御部 9 は、

外部バス 1 2 にて相互に接続されている。外部バス 1 2 は、内部バス 1 1 とは別の通信手段として作用する。CPU 1 は、システム制御部 3 を介して、外部バス 1 2 上の資源にアクセスする。

【 0 0 2 4 】

本実施の形態は、拡張可能な非線形画像編集装置である。この画像編集装置は、入力映像データとして、画像編集装置に既知の任意のフォーマットを受け取り、種々の編集コマンドに応じて映像の出力ストリームを生成することができる。

【 0 0 2 5 】

この画像編集装置は、動的に拡張可能なモジュールの組を有している。これらのモジュールは、特定のデータフォーマット又はデータフォーマットのクラスに対して、速度及び／又は画像品質を最適化するように処理することができる。かかるモジュールは、特定のフォーマット又はフォーマットのクラスに特有の特徴に基づく技術を用いている。この技術は、周波数領域と圧縮領域における処理技術を含むが、これに限定されない。

【 0 0 2 6 】

フォーマットモジュールは、標準的なビデオ編集 API に準拠している。モジュールは、システムに対して、追加され、削除され、又は置き換えられてもよい。

【 0 0 2 7 】

画像編集装置は、ビデオモジュールがデータに最適な処理を行うことができない場合には、ビデオデータを共通フォーマットに変換する。

【 0 0 2 8 】

モジュールは、ビデオデータをデータフォーマットのあるクラス中で処理する場合に、そのデータを中間フォームに変換してもよい。

【 0 0 2 9 】

映像データの本来のフォーマット中でビデオ処理操作を実行することは、今後、ウェーブレットやフラクタルに基づくフォーマットのような他のビデオフォーマットに対して有効であることがわかるであろう。

【 0 0 3 0 】

この画像編集装置は、拡張可能であるので、このような新しい展開が起きた場合には、対応するビデオモジュールをシステムに追加することによって、それらのフォーマットを利用することができる。

【 0 0 3 1 】

画像編集装置においては、ハードディスク 6 に蓄積された画像データに対して非線形の画像処理が行われる。この非線形編集について、図 2 に示す状態遷移図を参照して説明する。

【 0 0 3 2 】

画像データの非線形編集は、画像編集装置のハードウェアで実行されるオペレーティングシステム上で実現される。このオペレーティングシステムは、ファイルシステムとファイル操作に対するシステムレベルの操作、ファイル選択を含むグラフィカルユーザーインターフェース操作、ウィンドウ及びイベント処理、画像処理ファイルに対する標準操作を提供するメディアライブラリを提供する。

【 0 0 3 3 】

メディアライブラリは、画像ファイルを、開き、読み出し、書き込み、閉じ、同定することができるものでなければならない。

【 0 0 3 4 】

図 2 において、状態 S 0 1 では、画像編集装置において、編集コマンドが、グラフィカルユーザーインターフェースにより、又はバッチ処理により入力される。入力された編集コマンドを編集リストに記録する。

【 0 0 3 5 】

通常の編集コマンドとしては、入力ファイルをインポートし、インポートした画像ファイルに対する画像処理等がある。具体的には、編集コマンドとしては、カット内の始点と終点を指定し、タイトルや他の画像を重畳し、カットとカットとの間におけるワイプとディソルブのようなトランジション（遷移）を指定するためのものがある。

【 0 0 3 6 】

画像データに対する標準操作を提供するメディアファイルは、オペレーティン

グシステムによって提供されるファイルダイアログを用いて特定される。画像データ等のマルチメディアデータのコピー、カット、削除、移動、クリッピング等の編集操作は、ユーザーによってグラフィカルインターフェース上でマウスとメニューコマンドを用いて指定される。

【0037】

「出力生成」コマンドが与えられると、編集リストが生成される。また、その時にリストを作成してもよいが、ずっとコマンドを受けている時 (event loop: イベント ループ) に動的リスト作ってもよい。

【0038】

編集リストは、いくつかのトラックから構成され、各トラックは、画像、音声、効果のようなトラックタイプとクリップのリストを有している。クリップは、画像データ、音声データ又は効果データと、始点と期間の情報を含んでいる。

【0039】

ここで、編集リストの具体例として、2個の画像トラックと1個の効果トラックに対応するものについて、図3を参照して説明する。この編集リストは、C++を用いたクラス (class) として記述されたものである。しかし、編集リストとこれに伴う構成は、種々の他の同等な方法により表現することが可能である。

【0040】

図3の3Aに示すように、クラス (class) 編集リスト (EditList) は、絶対時間での開始フレームであるstartTimeと、フレーム単位での全ビデオの期間であるdurationと、Aトラック中のクリップの数であるaClipCountと、Aトラックの各クリップのクリップデータであるaSourcesと、Bトラック中のクリップの数であるbClipCountと、Bトラックの各クリップのクリップデータであるbSourcesと、効果トラックの効果の数であるeffectsCountと、効果データであるeffectsとを有している。

【0041】

図3の3Bに示すように、クラス (class) AVクリップ (AVclip) は、絶対フレームでの開始時間であるstartTimeと、このクリップのフレームの数であるdurationと、ビデオ圧縮フォーマットを表すmediaTypeと、この入力ストリームの

ファイル名であるfilenameと、ファイルの開始フレームであるframeOffsetとを有している。

【0042】

図3の3Cにおいては、クラス(class)効果(Effect)は、絶対フレームでの開始時刻であるstartTimeと、この効果のフレームの数であるdurationと、標準フォーマットの効果のタイプ(ワイプ等)を表すeffectTypeと、種々の効果パラメータ(スタート百分比、エンド百分比、方向等)であるeffectParamsとを有している。

【0043】

編集リストは、複数のトラックの間の関係を示すものであるので、時間軸に沿ってグラフィック表示することができる。図4は、編集リストとして、上述のように2個の画像トラックと1個の効果トラックに対応する具体例を示している。

【0044】

図4においては、図4の4Aには第1の画像トラックが、図4の4Bには効果トラックが、図4の4Cには第2の画像トラックが、図4の4Dには時間軸が表示されている。

【0045】

図4の4Aの第1の画像トラックには、時刻t0から時刻t3までの期間にわたって、第1のビデオクリップaが配置されている。

【0046】

図4の4Bの効果トラックにおいては、時刻t0から時刻t1までの期間にわたってタイトルマスクb1が、時刻t2からt3までの期間にわたってワイプトランジションb2が配置されている。

【0047】

図4の4Cの第2の画像トラックには、時刻t2から時刻t4までの期間にわたって第2のビデオクリップcが配置されている。

【0048】

図4の4Dには、時間の経過の方向に時刻t0, t1, t2, t3, t4の順序で表示した時間軸が表示されている。

【 0 0 4 9 】

このグラフィック表示においては、第 1 の画像トラックに含まれるビデオクリップ a と、第 2 の画像トラックに含まれるビデオクリップ c とは、画像出力を生成するために結合されている。すなわち、第 1 の画像トラックに含まれるビデオクリップ a の終わりが第 2 の画像トラックに含まれるビデオクリップ c の始まりとオーバーラップする部分に、ワイプトランジション b 2 が挿入されている。

【 0 0 5 0 】

また、ビデオクリップ a の始まりには、タイトルマスク b 1 が重畳された画像出力が得られる。

【 0 0 5 1 】

状態 S 0 2 においては、状態 S 0 1 において編集リストを編集セグメントに分解する処理が行われる。

【 0 0 5 2 】

上述したように、記録した編集リストは、時間軸方向に複数のトラックから構成される。この編集リストにおいて、編集トラック全体を時間軸に沿って分割した部分を編集セグメントという。したがって、編集リストは、時間方向に編集セグメントに分割することができる。

【 0 0 5 3 】

編集セグメントは、編集リストを構成する各映像トラックと各効果トラックが単一の値を取る最長の部分として定義される。したがって、図 4 のグラフィック表示に示した編集リストは、四つの編集セグメントに分割される。

【 0 0 5 4 】

分割された編集セグメントは、時刻 t 0 から時刻 t 1 までの期間の第 1 の編集セグメント、時刻 t 1 から時刻 t 2 までの期間の第 2 の編集セグメント、時刻 t 2 から時刻 t 3 までの期間の第 3 の編集セグメント、時刻 t 3 から時刻 t 4 までの期間の第 4 の編集セグメントになる。

【 0 0 5 5 】

ここで、任意の又は全てのトラックデータと効果トラックデータは、編集セグメント単位で空 (ϕ) となり得ることに注意する。

【 0 0 5 6 】

このように、図 4 に示した編集リストのグラフィック表示の例を編集セグメントに分割すると、図 5 に示すようになる。

【 0 0 5 7 】

すなわち、図 5 の 5 A に示す第 1 の編集セグメント (edit segment 1) は、時刻 t_0 から時刻 t_1 までに対応するもので、開始 (Start) t_0 、期間 (Duration) $t_1 - t_0$ 、第 1 のビデオ (Video 1) a 、第 2 のビデオ (Video 2) ϕ 、効果 (Effect) b_1 である。ここで、 ϕ は対象が存在しない「空」であることを示している。

【 0 0 5 8 】

図 5 の 5 B に示す第 2 の編集セグメント (edit segment 2) は、時刻 t_1 から時刻 t_2 までに対応するもので、開始 t_1 、期間 $t_2 - t_1$ 、第 1 のビデオ a 、第 2 のビデオ ϕ 、効果 ϕ である。

【 0 0 5 9 】

図 5 の 5 C に示す第 3 の編集セグメント (edit segment 3) は、時刻 t_2 から時刻 t_3 までに対応するもので、開始 t_2 、期間 $t_3 - t_2$ 、第 1 のビデオ a 、第 2 のビデオ c 、効果 b_2 である。

【 0 0 6 0 】

図 5 の 5 D に示す第 4 の編集セグメント (edit segment 4) は、時刻 t_3 から時刻 t_4 までに対応するもので、開始 t_3 、期間 $t_4 - t_3$ 、第 1 のビデオ ϕ 、第 2 のビデオ c 、効果 ϕ である。

【 0 0 6 1 】

編集セグメントが一旦構築されると、映像出力の生成を開始することができる。

【 0 0 6 2 】

本実施の形態は、状態 S_03 で行う処理において、映像出力の生成を、フォーマットや、フォーマットのクラスに対応させて行うものである。このような方法を、図 6 を参照して説明する。

【 0 0 6 3 】

図 6 の右側の S 1 4 は、対応したフォーマット内で映像出力を本実施の形態の編集方法である。図 6 の左側は、例えば 3 原色信号からなる RGB 信号や、輝度信号と二つの色差信号からなる Y C r C b 信号のようなベースバンド信号に復号してから映像出力を生成する従来の方法である。

【 0 0 6 4 】

ステップ S 1 4 においては、入力フォーマットが出力フォーマットと同一であるか判断する。入力フォーマットが出力フォーマットと同一であると判断されるとステップ S 1 6 に進み、同一ではないと判断されると従来の方法によるステップ S 1 0 に進む。

【 0 0 6 5 】

ステップ S 1 6 においては、出力フォーマットが入力フォーマットと同じであるので、画像編集装置はそのフォーマットに対応するビデオモジュール (VideoModule) を探し、そのビデオモジュールに効果処理を引き渡す。

【 0 0 6 6 】

対応するビデオモジュールの探索は、通常は容易な操作である。すなわち、全てのビデオモジュールのリストを探して、ビデオモジュールのフォーマットコードとビデオファイルに埋め込まれたオペレーティングシステムが提供する標準ビデオフォーマット (又は圧縮フォーマット) との比較を繰り返すことによって実現される。

【 0 0 6 7 】

ステップ S 1 7 においては、そのビデオモジュールにおける、効果を適用する関数である ApplyEffect を呼び出す。

【 0 0 6 8 】

ここで、ビデオモジュールは、画像処理用の標準 API (application program interface) を有する動的にロードされるライブラリである。図 7 はクラス (class) ビデオモジュール (VideoModule) のインターフェースである。

【 0 0 6 9 】

このビデオモジュールは、標準フォーマット特定子であるフォーマット (form

at) と、標準クラス特定子であるクラス (class) と、フレームを RGB に復号する ConvertToBaseband と、フレームをフォーマットに復号する (例えば DCT レベルに) ConvertToClass と、このフォーマットの知識を用いてトランジション効果を適用する ApplyEffect とを有している。

【 0 0 7 0 】

本実施の形態においては、フォーマットは、例えば、いわゆる DV 又はいわゆる MPEG-2 のような公開された標準フォーマットを表す整数である。クラスは、ビデオモジュールが属するフォーマットのグループを示す。例えば DV と motion JPEG は、共に「I-フレームのみの DCT に基づくビデオフォーマット」のクラスに属する。あるクラス内のビデオフォーマットはある中間フォームに変換することができる。この中間フォームは、速度、スペース及び／又は品質の点に関して圧縮されていないフォーマットよりも好ましい。したがって、DV と motion JPEG は、共に量子化されていない DCT 係数から構成されるフォーマットに変換することができる。

【 0 0 7 1 】

なお、従来では、効果を持たない編集セグメントに対しては、画像編集装置は、フレームに基づいたビデオフォーマットの全フレームの圧縮領域におけるコピーを行う。一方、効果を持つ編集セグメントに対しては、セグメント内の各フレームを画素領域に複合して画素領域で効果を適用され、その後にそのフレームが再圧縮されて出力ストリームに加えられる。ビデオ出力の全部が完成するまで、この処理が各編集セグメントについて続行される。

【 0 0 7 2 】

一方、従来の方法による出力映像の生成は、ステップ S 1 0 において行われる。

【 0 0 7 3 】

ステップ S 1 0 において、ステップ S 1 1 において入力フレームを通常の RGB フォーマットに復号し、ステップ S 1 2 において RGB 領域で効果を実行し、ステップ S 1 3 においてフレームを出力フォーマットに再符号化する。

【0074】

編集装置における各フレームのベースバンドへの復号は、符号化／復号器 (coder/decoder; codec) を用いることによって行われる。

【0075】

上述したような、フォーマットに対応して映像出力を生成する方法は、さらにフォーマットのクラスに対応して映像出力を生成するように拡張することができる。フォーマットのクラスを用いる一つの編集セグメントの処理は、図8に示すようになる。

【0076】

ステップS20においては、入力フォーマットを所望の出力フォーマットと比較する。これらが同一であればステップS21に進み、同一でなければステップS23に進む。

【0077】

入力フォーマットが出力フォーマットと同一であれば、図6におけるS16とS17で記載した同じ方法により処理を継続する。すなわち、ステップS21においては、ステップS16と同様に、対応するビデオモジュールを探索する。ステップS22においては、ステップS17と同様に、ビデオモジュールのApplyEffectを用いて効果を適用する。

【0078】

一方、ステップS23においては、入力クラスは出力クラスと比較される。クラスも異なるなら、図6に示したステップS10で前に示したように、入力フレームをベースバンド信号を用いて画素領域で結合する従来の方法に進む。

【0079】

クラスが同一であれば、ステップS24に進み、入力を中間フォームに変換する。

【0080】

上述したように、フォーマットのクラスは、ビデオモジュールが属するフォーマットのグループを示すものであり、例えばDVとmotion JPEGは、共に「I-フレームのみのDCTに基づくビデオフォーマット」のクラスに属する。

【 0 0 8 1 】

これらのフォーマットのクラスが共通する画像データは、共通フォーマットである中間フォームに変換される。

【 0 0 8 2 】

ステップ S 2 5 において、この中間フォーム上で、ビデオモジュール (VideoModule) の ApplyEffect によって処理が適用される。最後に、ステップ S 2 5 の結果は、ステップ S 2 6 において中間フォームから出力フォーマットに変換される。

【 0 0 8 3 】

続いて、図 9、図 1 0、図 1 1 を参照して、画像編集の処理のより詳細な例について説明する。

【 0 0 8 4 】

図 9 のように、異なったフォーマットの二つの入力ファイルがディソルブ (dissolve) トランジションを用いて結合され、いわゆる DV ファイルとして出力されるものとする。

【 0 0 8 5 】

第 1 の入力「ウェディング-DV. avi」はいわゆる DV フォーマットである。第 2 の入力「ピクニック-MJPEG. avi」はいわゆる motion JPEG フォーマットである。そして、これらを結合したいいわゆる DV フォーマットにより、「ハッピー・デー-DV. avi」を出力している。

【 0 0 8 6 】

図 1 0 は、いわゆる DV といわゆる MJPEG フォーマット、それらの共通のフォーマットのクラスである I - フレームのみの DCT フォーマットのクラスを識別するための、いくつかの便利な定数を簡単に列記している。

【 0 0 8 7 】

これらの定数は、FourCC(a,b,c,d)によって、24ビットシフトされた第1の数 a、16ビットシフトされた第2の数 b、8ビットシフトされた第3の数 c、及び第4の数 d の論理和として定義されている。

【0088】

すなわち、FORMAT_DVは、FourCC('D','V','C','S')と、FORMAT_MJPEGは、FourCC('M','J','P','G')と、FORMAT_IDCTは、FourCC('I','D','C','T')と、それぞれ定義される。

【0089】

このようなフォーマットのクラスを用いた、図9に示したような入力ファイルの結合について、図11を参照して説明する。

【0090】

ステップS30においては、異なるフォーマットであるが同じクラス（IフレームDCT）の二つのファイル「ウェディング-DV. avi」と「ピクニック-MPEG. avi」が与えられる。これらのファイルは、以下のステップにより、ディソルブ（dissolve）トランジションにて結合される。

【0091】

ステップS31において、二つの入力ファイルのフォーマットを決定する。フォーマットはオペレーティングシステムのメディアライブラリを用いて決定される。これにより、「ウェディング-DV. avi」と「ピクニック-MPEG. avi」のフォーマットは、それぞれ「FORMAT_DV」と「FORMAT_MJPEG」と決定される。

【0092】

これらのフォーマットは同一ではないので、ステップS32に進み、それぞれのフォーマットに対するビデオモジュールを探索する。ビデオモジュールの探索の結果、「ウェディング-DV. avi」と「ピクニック-MJPEG. avi」のビデオモジュールは、それぞれ「DVVideoModule」と「MJPEGVideoModule」となる。

【0093】

ステップS33において、フォーマットが等しくないので、ビデオモジュールにそのクラスを照会し、これらを相互に比較する。すなわち、「ウェディング-DV. avi」と「ピクニック-MJPEG. avi」にそれぞれ対応するクラスモジュール「DVVideoModule」と「MJPEGVideoModule」からクラスを取得する。

取得したクラスは、同じく「CLASS_IDCT」となる。

【0094】

ステップS34においては、ステップS33で求めたクラスが等しいので、入力ビデオモジュール上でConvertToClass（）を呼び出すことにより、各入力ビデオを共通クラスフォーマットに変換する。

【0095】

ステップS35において、IDCTVideoModuleが、変換されたDCT形式のデータをディソルブ効果を用いて結合する。

【0096】

ステップS36において、中間フォームは出力フォーマット（DV）に変換される。

【0097】

最後にステップS37において、編集セグメントが目的ファイルに出力される。この処理は従来のビデオ画像編集装置が効果を実行する前にビデオを画素フォーマットに復号する方法と同様である。

【0098】

但し、本発明においては、フレームが部分的に復号されるのみであることが異なる。この部分的復号は、完全なソフトウェア復号よりも通常は迅速に行うことが可能であり、上述のようにフォーマットモジュールのクラスがフォーマット特有の最適化のクラスを利用することを許容する。

【0099】

このように、本実施の形態においては、ビデオモジュールを用いることにより、フォーマットに固有の知識をカプセル化し、速度及び／又は画質について最適化された編集操作を実行する。

【0100】

本実施の形態においては、ビデオモジュールを、ハードウェアディスクの特定ディレクトリに配置された動的にリンクされるライブラリとして実現している。

【0101】

画像編集装置においては、編集装置は、複数のビデオモジュールにアクセスし

て、ビデオモジュールを更新するための手段を有している。ビデオモジュールは、単一の編集セッション、又は編集セッションの間に、動的に更新することができる。すなわち、動的リンクライブラリファイルを追加、削除、置き換えることにより、画像編集装置が利用できるフォーマットは自動的に更新される。

【0 1 0 2】

次に、DCTに基づくビデオを用いる効果の処理について説明する。ビデオモジュールを用いることにより、本実施の形態は、任意の数の圧縮フォーマットも効果的に支持する。DCTに基づいて行うデジタルビデオ圧縮の利点について述べる。

【0 1 0 3】

DCTに基づくビデオ圧縮方法は、各フレームを規則的なサイズの（しばしば 8×8 ）画素ブロックに分割する。これらのブロックは画素コンポーネント（例えばY, Cr, Cb）によってマクロブロックに構成される。マクロブロックの2つの例を図12に示す。

【0 1 0 4】

図12においては、 8×8 のブロックを単位として、4 : 1 : 1マクロブロック22と4 : 2 : 0マクロブロック23の両方が6つのブロックから構成される。

【0 1 0 5】

すなわち、4 : 1 : 1マクロブロック22は、4個のYブロック22a、1個のCrブロック22b及び1個のCbブロック22cから構成される。また、4 : 2 : 0マクロブロック23は、4個のYブロック23a、1個のCrブロック23b及び1個のCbブロック23cから構成される。このように、Yブロックと、CrブロックとCbブロックが重なってマクロブロックになる。

【0 1 0 6】

図12では2つの例のマクロブロック構造が示されているが、多くの同様なマクロブロック構造が通常使用されている。

【0 1 0 7】

図13には、ビデオのフレームをマクロブロックに分割する2つの例を示して

いる。

【0 1 0 8】

すなわち、図 1 3 においては、ビデオのフレーム 3 3 は、縦横に 1 4 × 5 の 4 : 1 : 1 マクロブロック 3 1 から構成されている。4 : 1 : 1 マクロブロック 3 1 は、Y 信号に対応する 4 個の Y ブロック 3 1 a と、C r 信号に対応する第 5 のブロック 3 1 b と、C b 信号に対応する第 6 のブロック 3 1 c とから構成されている。

【0 1 0 9】

また、ビデオのフレーム 3 4 は、縦横に 7 × 1 0 の 4 : 2 : 0 マクロブロック 3 2 から構成されている。4 : 2 : 0 マクロブロック 3 2 は、Y 信号に対応する 4 個の Y ブロック 3 2 a と、C r 信号に対応する第 5 のマクロブロック 3 2 b と、C b 信号に対応する第 6 のブロック 3 2 c とから構成されている。

【0 1 1 0】

続いて、トランジションの一例としてワイプトランジションについて、図 1 4 を参照して説明する。

【0 1 1 1】

図 1 4 に示すワイプトランジションは、文字「B」を表示する第 1 の入力フレーム 4 1 と文字「A」を表示する第 2 の入力フレーム 4 2 とを用いて、出力フレーム 4 3 に対して、片方から第 1 の入力フレーム 4 1 と第 2 の入力フレーム 4 2 との一方を拭き去るように消していき、後から他方のフレームを現していくものである。

【0 1 1 2】

出力フレーム 4 3 においては、第 1 の入力フレーム 4 1 の一部 4 1 a と、第 2 の入力フレーム 4 2 の一部 4 2 b とが表示されている。

【0 1 1 3】

出力フレームのトランジションエッジ 4 4 に沿ったマクロブロック 4 5 は、第 1 の入力フレーム 4 1 と第 2 の入力フレーム 4 2 とから構成されている。これらの 2 つの入力は、可能なら周波数領域で結合されるが、周波数領域で結合することができないときには必要なブロック数だけ画素領域で結合される。

【0 1 1 4】

ここで、ワイプ位置を跨がないマクロブロックは、周波数領域でコピーにより処理できる。ワイプ位置を跨ぐ場合にも周波数領域で処理できることが多い。このことについては、後述する。

【0 1 1 5】

例えば、三次元効果や変形作画 (warping) のような非マクロブロック局所効果を伴うさらに複雑な効果は、周波数領域で実現することが困難なので、従来方法 (ベースバンド) で取り扱われる。

【0 1 1 6】

このワイプトランジションは、多くの場合、入力フレームから圧縮マクロブロックをコピーすることにより実行することができる。トランジションエッジにまたがる少数のマクロブロックは、入力フレームから適当なブロックを D C T 係数レベルに部分的に復号し、それらを予め計算されたワイプ行列を用いて結合し、そして最終的なビット・ストリームに再圧縮することで生成される。ある場合には、コピーされたブロックと結合されたブロックは、一つのビットストリームに許容されたビット数以内に収まらないので再量子化が行われる。

【0 1 1 7】

アイリス効果及びビットマスクオーバーレイのような他の「局所」効果も、大半のマクロブロックをソースから直接にコピーし、残りのマクロブロックについて D C T 領域処理技術を用いて、同様の方法で実行することができる。

【0 1 1 8】

クロスディソルブは入力マクロブロックをコピーすることでは達成できない。なぜなら、それは全てのマクロブロック入力ブロックの線形結合であるからである。しかし、D C T の特有の性質は、クロスディソルブが入力 D C T 係数の重み付けされた結合を用いて簡単に計算できる。

【0 1 1 9】

出力 D C T 係数 $c_{i,j}$ は、次式に示すように計算することができる。

$$c_{i,j} = \alpha \times a_{i,j} + (1 - \alpha) \times b_{i,j}$$

ここで、 $c_{i,j}$ は、量子化されていない D C T の行 i 、列 j の出力係数である。

$a_{i,j}$ は、量子化されていない DCT の行 i 、列 j の入力 A の係数である。 $b_{i,j}$ は、量子化されていない DCT の行 i 、列 j の入力 B の係数である。 α は、入力 A が受けるディソルブの割合であり、 $0 \leq \alpha \leq 1$ 、 0 の範囲にある。

【0 1 2 0】

なお、DCT 領域での処理技術の優れた概論として、カリフォルニア大学バークレー校の「電気技術及びコンピュータ科学」誌に掲載された、シーフー・チャン (Shih-Fu Chang) 氏の 1993 年の工学博士号論文「マルチメディア・ネットワーク・ビデオ・サービス用のビデオ信号の合成と操作 (Compositing and Manipulation of Video Signals for Multimedia Network Video Services)」がある。

【0 1 2 1】

さらに、本発明は、編集装置ハードディスク 6 に格納されたプログラムとしても実現され、同じくハードディスク 6 に格納されたメディアファイルを使用する。複数のハードディスク又はその他の記憶用メディアの使用も、本発明の範囲に入る。

【0 1 2 2】

例えば、図 1 に示した編集装置におけるハードディスクには、上述した画像編集を行うプログラムが記録される。このプログラムは、メモリ 4 に読み込まれ、オペレーティングシステム上で実行され、上述したような画像編集装置における編集を実現する。

【0 1 2 3】

このプログラムは、入力画像信号の形式を特定し、上記入力画像信号に対する編集操作を入力され、上記入力画像信号を入力された編集操作にしたがって編集し、編集された上記入力画像信号を出力ストリームに変換し、特定した上記入力画像信号のフォーマットやフォーマットのクラスに応じて、当該形式に対応する編集を行うように制御するものである。上記入力画像信号を中間フォームに変換しても良い。

【0 1 2 4】

なお、上述の実施の形態においては、フォーマットとして DV や motion JPEG

G等を、フォーマットのクラスとしてはI-フレームを例示したが、本発明はこれに限定されない。他のフォーマットやフォーマットのクラスを適用することができるのはいうまでもない。

【 0 1 2 5 】

【発明の効果】

上述のように、本発明によると、種々のフォーマットの圧縮画像信号を伸長することなく編集することができる。したがって、本発明によると、画像信号の編集を高速に、品質を劣化させることなく行うことができる。

【図面の簡単な説明】

【図 1】

映像編集装置の全体の構成を概略的に示すブロック図である。

【図 2】

映像編集装置における編集処理の手順を示す状態遷移図である。

【図 3】

編集リスト、AVクリップ及び効果の構造体を示す図である。

【図 4】

簡単なビデオ編集の時系列を示す図である。

【図 5】

簡単な編集リストを分割した編集セグメントを示す図である。

【図 6】

編集セグメントの処理を示すフローチャートである。

【図 7】

クラスビデオモジュールを示す図である。

【図 8】

編集セグメントの処理を示すフローチャートである。

【図 9】

異なったフォーマットを有する2個の入力ファイルからなる編集セグメントを示す図である。

【図 1 0】

クラスは同じであるがフォーマットが異なる入力ビデオを用いる編集セグメントを処理する例を示す図である。

【図 1 1】

フォーマットとクラスの定義を示す図である。

【図 1 2】

典型的なマクロブロックの構造を示す図である。

【図 1 3】

マクロブロックの画面への配置を示す図である。

【図 1 4】

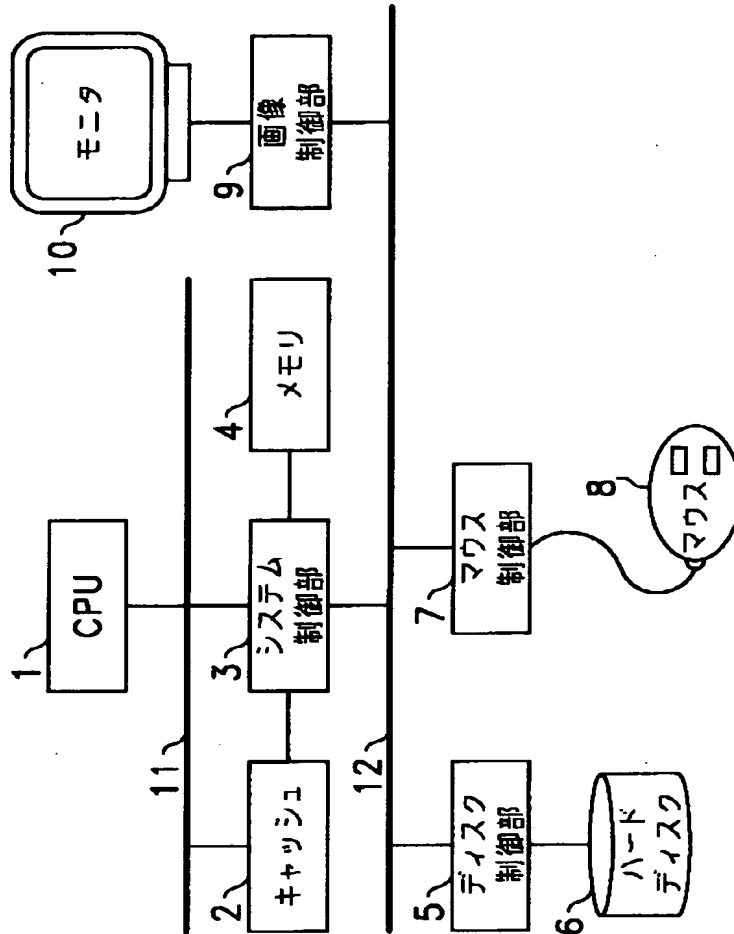
大部分がマクロブロックの複写で構成されるワイプトランジションを示す図である。

【符号の説明】

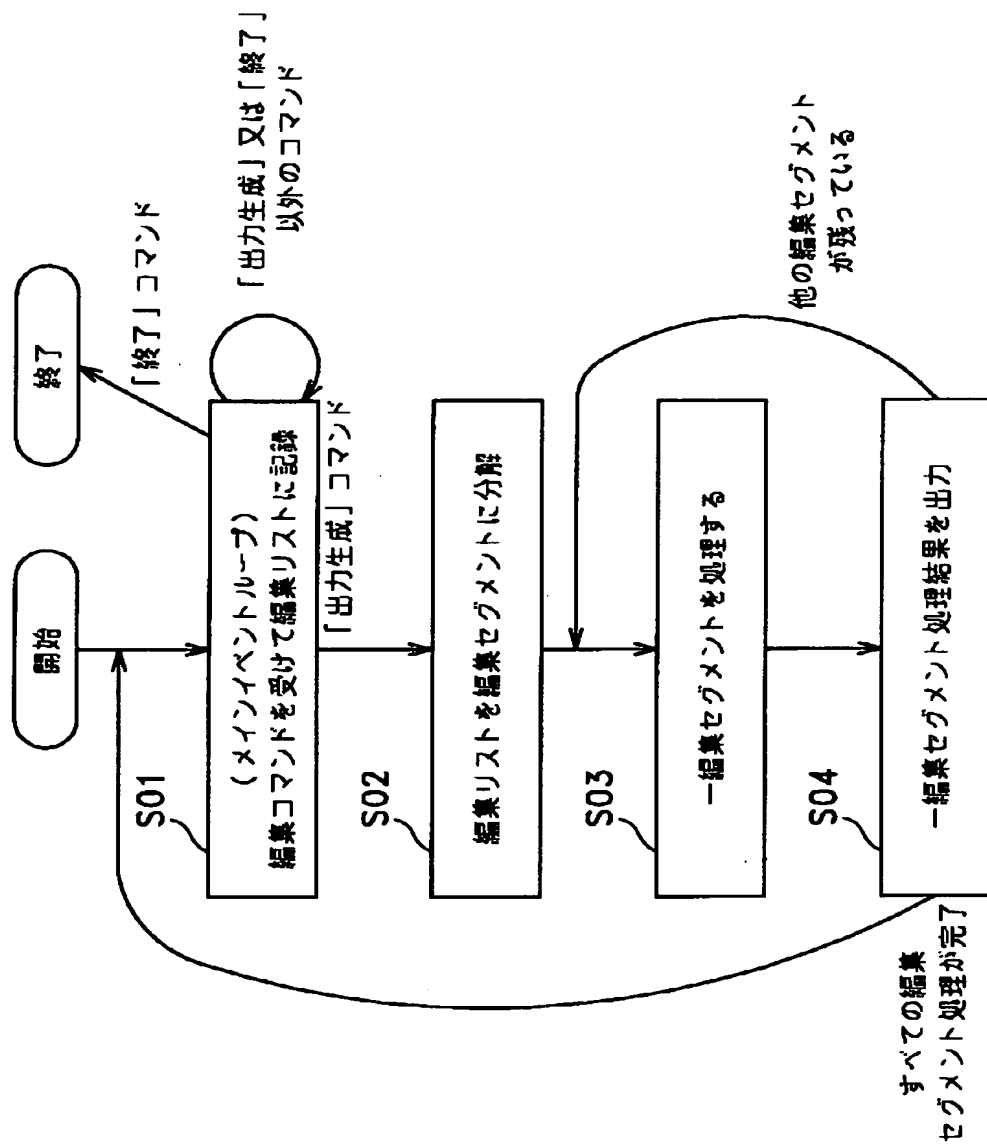
1 CPU、2 キャッシュ、3 システム制御部、4 メモリ、5 ディスク制御部、6 ハードディスク、7 マウス制御部、8 マウス、9 ビデオ制御部、10 モニタ

【書類名】 図面

【図 1】



【図 2】



【図 3】

3A

```
class EditList {
    int    startTime;    /* 開始時刻 (絶対フレーム)          */
    int    duration;     /* 全ビデオの期間 (フレーム)        */
    int    aClipCounty;  /* Aトラック中のクリップの数        */
    AVClip aSources[];   /* Aトラックの各クリップのクリップデータ */
    int    bClipCount;   /* Bトラック中のクリップの数        */
    AVClip bSources[];   /* Bトラックの各クリップのクリップデータ */
    int    effectsCount; /* 効果トラックの効果の数          */
    Effect effects[];    /* 効果データ                        */
};
```

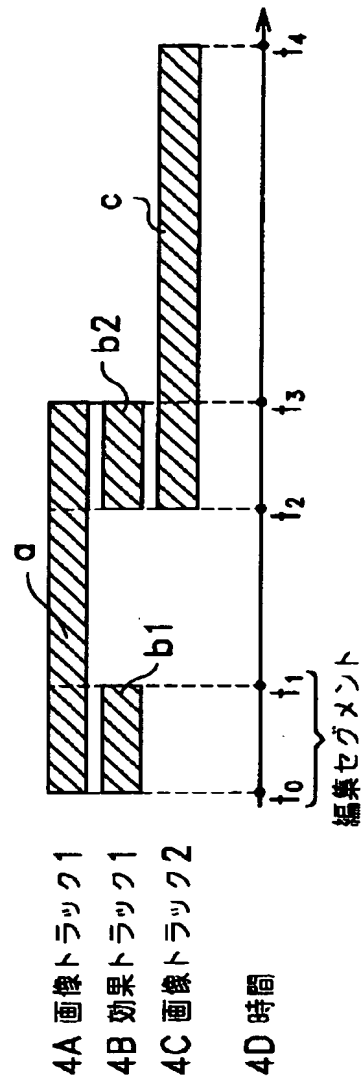
3B

```
class AVClip {
    int    startTime;    /* 開始時刻 (絶対フレーム)          */
    int    duration;     /* このクリップのフレームの数        */
    int    mediaType;    /* ビデオ圧縮フォーマット (標準)      */
    char   filename[];   /* この入力ストリームのファイル名    */
    int    frameOffset;  /* ファイルの開始フレーム            */
};
```

3C

```
class Effect {
    int    startTime;    /* 開始時刻 (絶対フレーム)          */
    int    duration;     /* この効果のフレームの数            */
    int    effectType;   /* 標準フォーマットの効果のタイプ (ワイプ等) */
    void   *effectParams; /* 種々の効果パラメータ (スタート百分比、
                                /* エンド百分比、方向等)          */
};
```

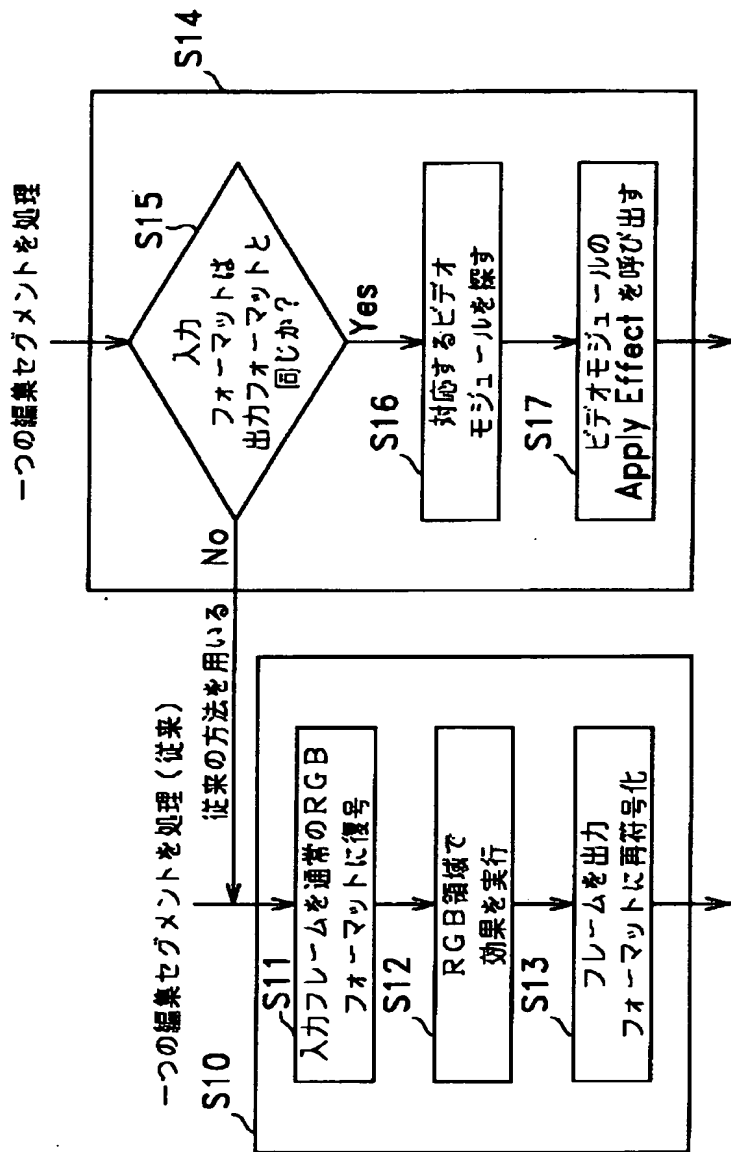
【図 4】



【図 5】

5A	編集セグメント 1	
	開始	t_0
	期間	$t_1 - t_0$
	ビデオ 1	a
	ビデオ 2	\emptyset
	効果	b1
5B	編集セグメント 2	
	開始	t_1
	期間	$t_2 - t_1$
	ビデオ 1	a
	ビデオ 2	\emptyset
	効果	\emptyset
5C	編集セグメント 3	
	開始	t_2
	期間	$t_3 - t_2$
	ビデオ 1	a
	ビデオ 2	c
	効果	b2
5D	編集セグメント 4	
	開始	t_3
	期間	$t_4 - t_3$
	ビデオ 1	\emptyset
	ビデオ 2	c
	効果	\emptyset

【図 6】



【図 7】

```
class VideoModel {
    int format;          /* 標準フォーマット特定子 */
    int class;           /* 標準クラス特定子 */

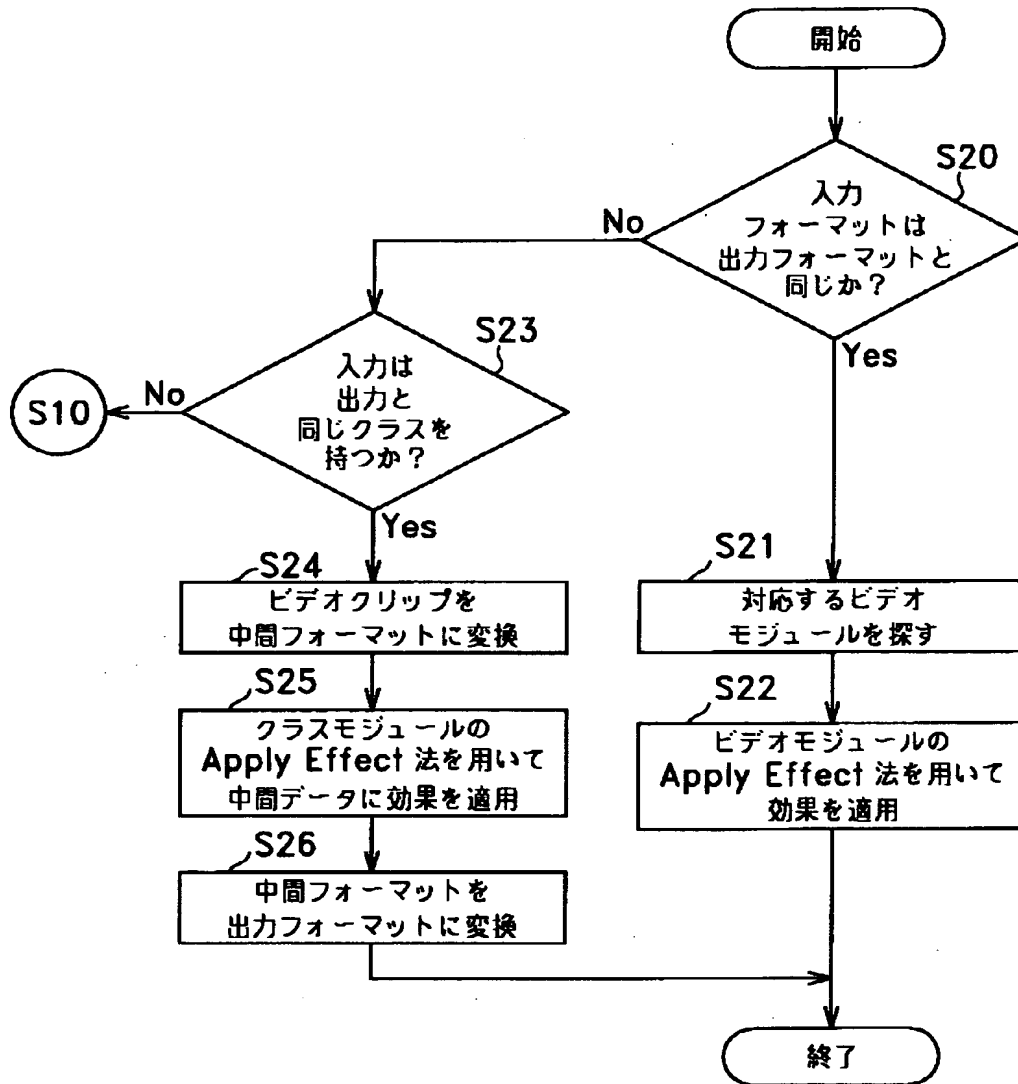
    // フレームを RGB に復号
    char *ConvertToBaseband(AVSource *pSource, int frame);

    // フレームをフォーマットに復号 (例えば DCT レベルに)
    void *ConvertToClass(AVSource * pSource);

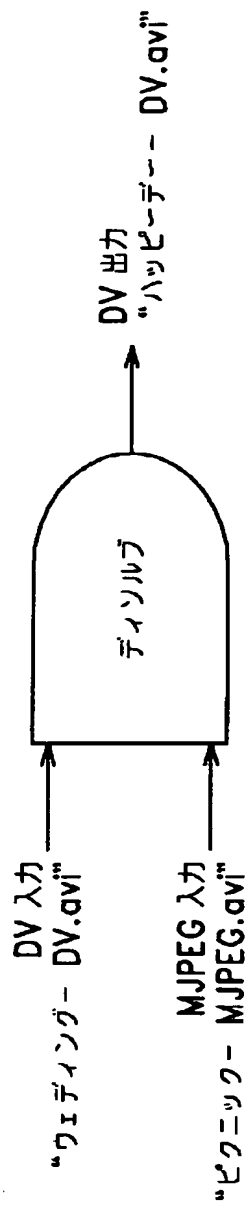
    // フォーマットのクラス (例えば DCT) からこのフォーマットに変換
    void *ConvertFromClass(Void * pClassData);

    // このフォーマットの特定の知識を用いてトランジション効果を適用
    int ApplyEffect(AVSource *pInputA, AVSource *pInputB,
                   AVSource *pOutputC, Effect *pEffect);
};
```

【図 8】



【図 9】



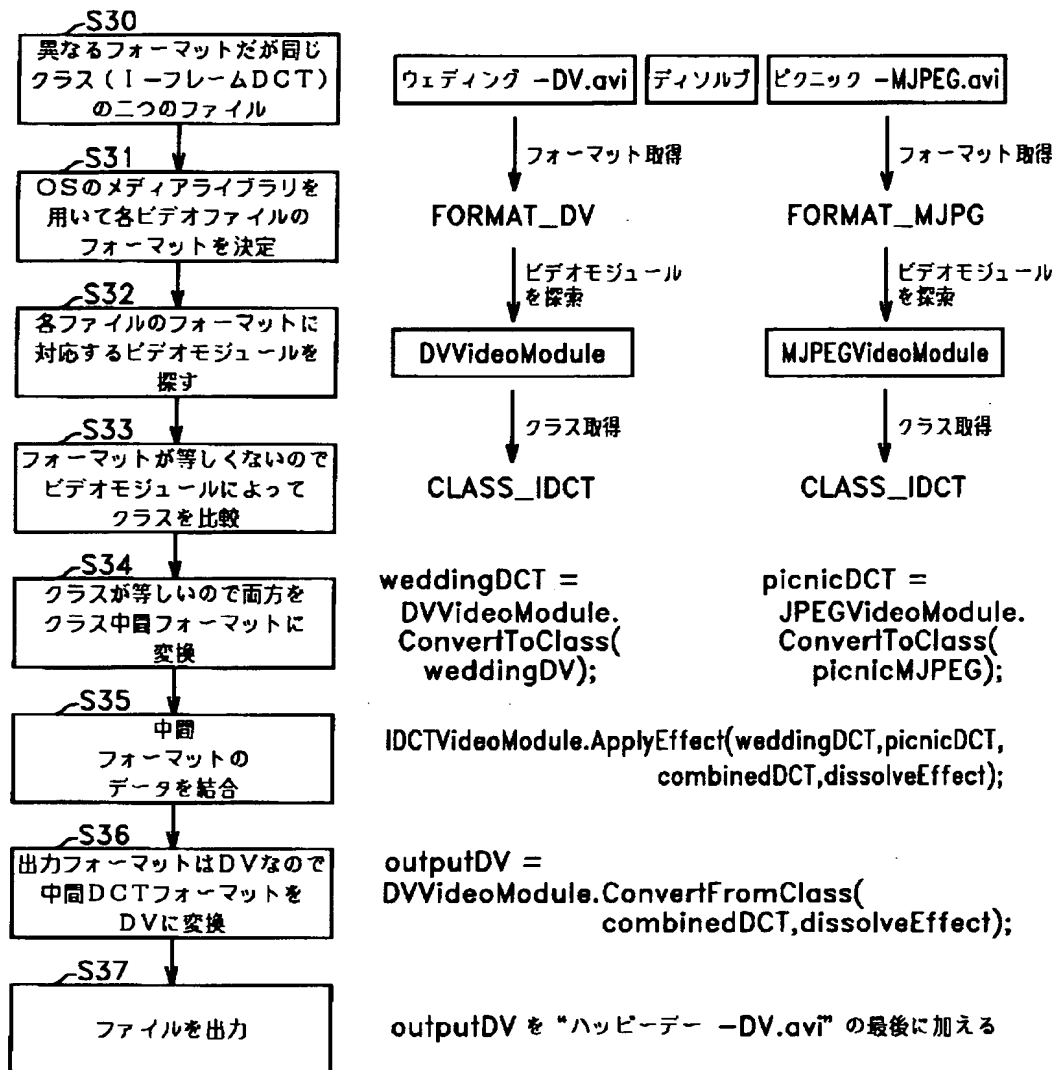
【図 1 0】

```
#define FourCC(a,b,c,d) ((a<<24) | (b<<16) | (c<<8) | (d))

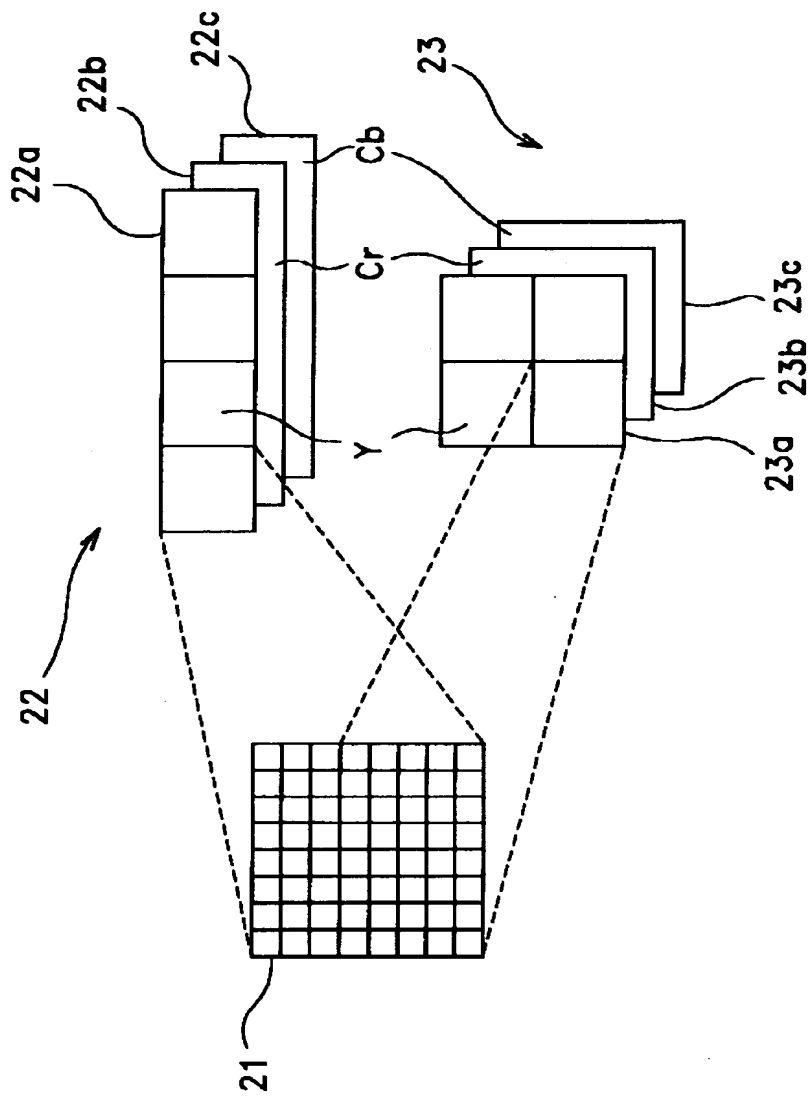
#define FORMAT_DV      FourCC('D','V','C','S')
#define FORMAT_MJPG    FourCC('M','J','P','G')

#define CLASS_IDCT      FourCC('I','D','C','T')
```

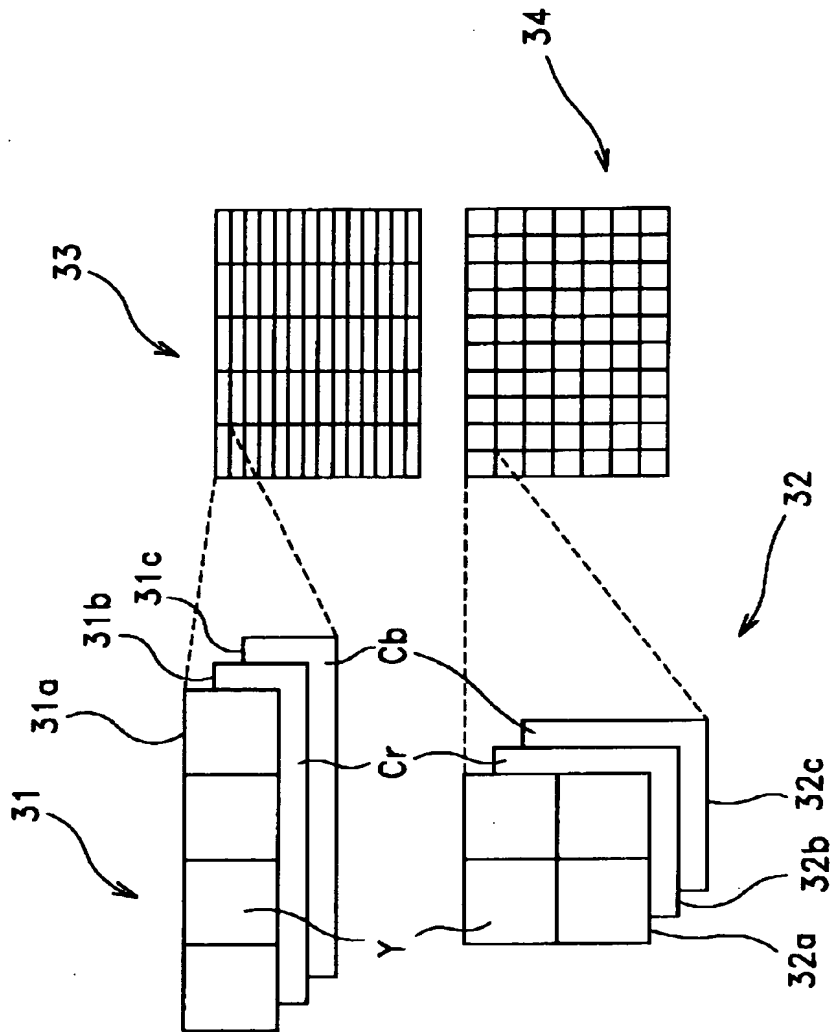
【図 1 1】



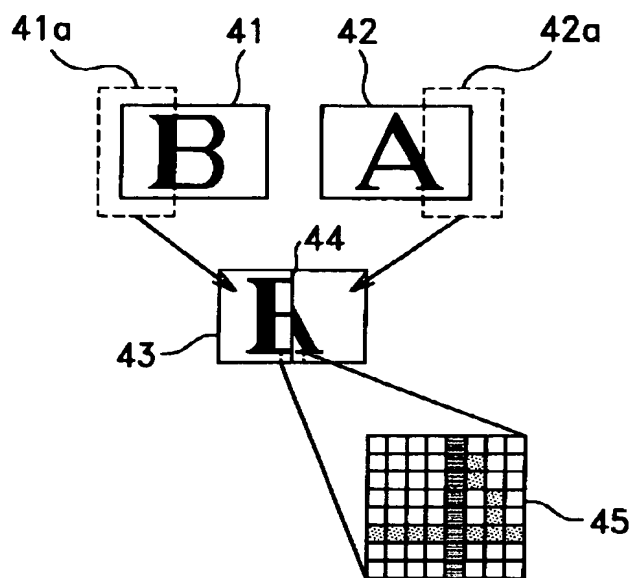
【図 1 2】



【図 1 3】



【図 1 4】



【書類名】 要約書

【要約】

【課題】 画像信号の編集を高速に、画質を劣化させないように行う。

【解決手段】 状態 S 0 1 においては入力コマンドが編集リストに記録され、状態 S 0 2 においては編集リストが編集セグメントに分解され、状態 S 0 2 においては一編集セグメントについて当該画像信号の形式に応じた処理が施され、状態 S 0 4 においては一編集セグメントの処理結果が出力される。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000002185]

1. 変更年月日	1990年 8月30日
[変更理由]	新規登録
住 所	東京都品川区北品川6丁目7番35号
氏 名	ソニー株式会社